# 关于元分析（Meta analysis）R程序包的综述（译）

原文：Polanin, J. R., Hennessy, E. A., & Tanner-Smith, E. E. (2017). A Review of Meta-Analysis Packages in R. Journal of Educational and Behavioral Statistics, 42(2): 206-242.

翻译：郭丽萍1，乔陆1

校对：范会勇1，胡传鹏2

(1, 渤海大学教育与体育学院, 锦州 121013;

2, Deutsches Resilienz Zentrum (DRZ), University Medical Center of the Johannes

Gutenberg University, 55131 Mainz, Germany)

**摘 要** 元分析（meta analysis）是一种分析多篇原始研究（primary）的综合效应值的统计技术。为了估计（estimate）元分析的效应模型，开源（open-source）统计环境（statistics environment）R迅速成为一个受欢迎的选择。元分析社区通过发展多种特定的元分析程序包从而推动了R程序包使用频率的增长。本研究的目的在于找到全部公开的、可获取的R程序包。通过全面的网络搜索，我们搜索到了63个程序包。为了帮助阐明它们在元分析领域的功能，我们对每一类R程序包都进行了描述，同时还为对R程序包感兴趣的研究者推荐了一些应用，提供了两种元分析程序包的简明指导，也为将来R程序包的创建者提供了一些建议。

**关键词**：元分析；效应值；统计软件；R

## 引言

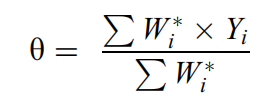
元分析是一种统计技术，这种技术通过综合大量研究的效应值，最终形成一个有意义的估计。与单个原始研究的结果相比，元分析研究提高了结果的普适性、精确度以及跨研究探索异质性的能力(Borenstein, Hedges, Higgins, & Rothstein, 2010; Pigott, 2012)。事实证明，元分析对政策制定者和实践者非常有用，因为元分析的结果为模棱两可的问题以及大量的研究文献提供了一个答案。因此，在过去30年里，元分析的发表数量呈指数增长(Willianms, 2012)。

传统的元分析是通过简单的方差倒数加权平均（inverse-variance-weighted）来估计综合效应值，并没有要求使用复杂的软件或高级计算能力（advance computing power）（Hedges & Olkin，1985）。分析师可以手动或通过普通的电子表格软件相对容易地计算加权平均值。然而，除了最简单的元分析任务外，使用统计软件和预编程序应用是更有效。尤其是在迭代或基于仿真的技术相关的研究中，这种封闭式计算是不可行，或者可能会引入偏差。

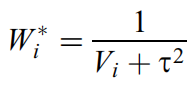
这种软件的一个快速增长的市场就是R统计环境。R是一种程序语言，是一种通过复杂的R网络中管理和维护的开源软件（R Core Team，2015）。它已成为许多领域的研究者和数据科学家们的主要选择（Rexer Analytics，2013）。在R中建立元分析软件包也是跟随这种趋势。ChenandPeace’s (2013) 和Schwarzer, Carpenter, and Rücker (2015b)的教科书中描述了许多基于R元分析的程序包，Neupane, Richer, Bonner, Kibret, and Beyene (2014)详细说明了可用于进行网络元分析的R程序包。然而，这些著作中并没有对所有的R程序包进行综合分析。因此，本研究的目的在于阐明可供研究者使用的R元分析程序包，解释两个热门元分析程序包的功能，并为程序包未来的改进提供一些建议。

## 背景

传统的元分析模型通过效应值大小的变异的倒数对效应值进行加权（例如，标准误平方） 来估计平均效应值。这样做会使那些有着大样本量，小变异的研究比那些有着大变异的研究对平均效应量的影响更大，这就是以往提到的固定效应模型（Borenstein et al 2010）。另一种流行的方法是估计随机效应模型，该模型除了通过效应值大小的变异的倒数进行加权外，还考虑到研究之间的差异。可以用以下公式进行表示：



θ表示元分析的平均效应值，Yi表示第i个研究的效应值，Wi\*表示第i个研究的随机效应权重。在随机效应模型中，权重用以下公式表示：



Vi第i个研究的效应值的变异（标准误平方），τ2表示研究间方差估计的平方。与固定效应模型相比，随机效应模型预测将会有更大的置信区间，并且每个研究的权重也会变得相似。在估计τ值时，虽然可以使用封闭的计算形式来解决 (DerSimonian & Laird, 1986)，但使用迭代最大似然无偏估计会减少偏倚（Veroniki et al., 2016）。热门的元分析R程序包，比如metafor (Viechtbauer, 2015) 或metaSEM (Cheung, 2015) 能够使用无偏估计计算出变异内容。

最近有关元分析方法的进展已经超过了传统的元分析模型，例如，单项研究在同一结果领域内报告多个效应量是很常见的，而在传统元分析假设下，通常每篇研究只会呈现一个效应量，比如研究者根据之前的规则选择一个效应量或简单的对某一研究中的所有效应量进行平均。除此之外，研究者可以用稳健方差估计把所有效应量整合进一个模型当中，充分考虑到研究内协方差对研究结果的影响。虽然之前的多元元分析方法曾提到过（e.g., Gasparrini, 2015），但并不是每个研究的稳健方差估计都需要知道影响效应值大小的研究内协方差分析结构。 Tanner-Smith, Tipton, and Polanin (2016)根据效应值大小之间的相关性提出了两种加权效应值的方法，robumeta软件包（Fisher & Tipton, 2015）也可以估计稳健方差元分析模型。

相似的进展也出现在各种常见的应用问题上。在利用元分析相关矩阵进行结构方程建模上，Cheung（2015）提出了两步建模方法。第一步，综合完整的或近似完整的相关矩阵；第二步，利用第一步中已综合的相关矩阵构建常用的结构方程模型，例如因子分析或潜变量模型。后来Wilson, Polanin和Lipsey等人进一步扩展了这个模型，使它能够适应更复杂的具有多水平相关性的数据集，Cheung(2015)的metaSEM也可以进行元分析结构方程模型建模。另一个进展是通过网络元分析（network Meta-analysis）来使用网络模型，其目的在于同时比较多种干预措施。利用两个研究可以呈现一个简单的网络模型，在研究1中，作者比较了接受干预的A组和没有接受干预的对照组B组；在研究2中，作者比较了研究1中接受干预的A组和另一个接受干预的C组，分析员可以使用网络分析作三个比较（A、B、C），即使在初始研究中这些组并没有进行直接比较。虽然网络元分析在社会科学和教育科学领域应用较少，但在医学领域却很普遍 (Neupane, Richer, Bonner,Kibret, & Beyene, 2014)。

近年来元分析在许多方面取得了很大的进展。接下来我们将通过讨论元分析R程序包来呈现这些进展。首先，我们对如何找到程序包，如何决定它们是否符合标准，如何编码它们的功能等问题提供了一个大致的概览。为了方便理解，我们根据程序包的功能进行了分类，并讨论不同类别的程序包，也为以后的程序员提供了建议。最后我们会对普遍使用的metafor程序包和robumeta程序包作简要指南。

## 方法

为了定位与检索到所有的有元分析功能的R程序包，我们进行了系统且全面的搜索。首先我们检索了warehouse R packages、the Comprehensive R Archive Network website(CRAN)和Revolution Analytics这三个网站，使用的检索词为：meta analysis、meta-analysis、systematic review。除此之外，我们还检索了Git Hub和Google Scholar，并向元分析领域的专家和开发R程序包的程序员请教。搜索结束后，每一个R程序包只呈现了一个作者。不管其整体的关注点、功能或局限，只要该程序包包括任何元分析的功能，都将包含在内。例如，CAMAN (Schlattmann, Hoehne, & Verba, 2015) 程序包最初主要用来产生有限混合模型，而该程序包也可以对P值进行基本的元分析，因此该程序包也被纳入到搜索结果中。

这个过程的目的是总结整个元分析包中可用的功能，并确定应用程序中的任何缺口，以便未来的程序员可以改进这些问题。

接下来，我们从程序包中编码和提取了特征（characteristics）。这一过程的目的是总结所有元分析程序包中可用的功能，并说明程序包中存在的缺点，以便于未来程序员可以改进这些问题。我们还为初学者提供了不同程序包的功能说明，因为对初学者来说知道哪些程序包可以提供哪些功能是最重要的。因此，我们编码了一个广泛的程序包特征（可能并不适用于所有的程序包），此外，我们还说明了程序包之间的灵活性和多样性。我们收集的信息包括程序包计算效应值或统计功效的功能、程序包如何处理遗失的信息、如何估计固定效应模型或随机效应模型、程序包是否适用于相关效应值、是否能够进行调节变量分析或敏感性分析（sensitive analysis）、是否能够评价出版偏倚、是否有绘图性能（plotting capacity）。我们还对不同程序包的主要功能进行命名，便于初学者可以迅速找到相应的程序包。一位作者进行了所有的编码，并与研究团队的其他成员定期举行编码会议，以讨论不同作者之间报告程序包功能的方式之间的细微差别。

编码完成后，我们对程序包进行了几次描述性分析。首先，我们对样本进行描述性分析，总结出整个程序包的主要功能，并根据不同的功能将程序包分为9个类别，分别是（1）general meta-analysis packages（2）genome（3）multivariate（4）diagnostic（5）specific（6）network（7）assessment of bias （8）Bayesian（9）graphical user interface（GUI）。其中七类反映了程序包的功能和特征，比如，general meta-analysis程序包执行基本的元分析功能，并覆盖传统元分析的所有功能。另外两类“genome”和“specific”代表了两类不同的程序包。genome程序包只具有基因类元分析功能，这是一个特殊的元分析领域，所以这些程序包一般不用于社会科学和医学这些常用的元分析中。specific程序包主要处理一些目标具有针对性的元分析程序。虽然这些类别是人为主观进行区分的，但它们仍具有启发性和实用性。最后我们使用R程序包计算了每一个程序包的日下载量和月下载量。

对R程序包的首次搜索结束于2015年1月30日，列表更新至2015年7月2日，编码过程结束于2015年9月29日。我们把所有程序包的功能编码在Excel表格中,并使用ggplot2程序包在R程序包中的创建摘要和图表(Wickham, 2009)。我们对编码时使用的R包特定的版本号进行了编码，尽管在编码结束之后可能还有其他程序更新。目前我们使用的所有R程序包为3.1.2版本。

搜索过程中共产生106个程序包。由于重复（n=32）、程序包不支持下载（e.g., RcmdrPlugin.Mac; n= 4）或不符合纳入标准（n=2），我们删除了38个程序包，另外，通过与专业同事的讨论以及之后在会议中呈现的初始研究发现 (Polanin& Hennessy, 2015)，又增加了5个程序包，因此，本研究纳入的程序包数量为63。

程序包最早创建于1999年，2012年创建程序包的数量最多。如表1（见附表）所示，这些程序包的功能大同小异，大多数程序包都可以计算效应量（66.7%）、计算固定效应模型（60.3%）或随机效应模型（74.6%），74.6%的程序包都可以制作元分析图表，但多数都依赖于其他少数几个程序包对图表进行渲染。相反，只有少数程序包可以处理缺失值（20.6%）、进行调节变量分析（31.8%）、处理相关效应值（44.4%）以及评价出版偏倚（22.2%）。最后，只有3.17%的程序包可以进行统计功效分析。图2呈现了不同类别的程序包。

表1程序包的特点

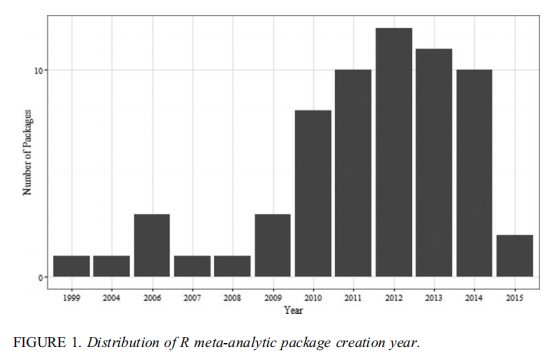
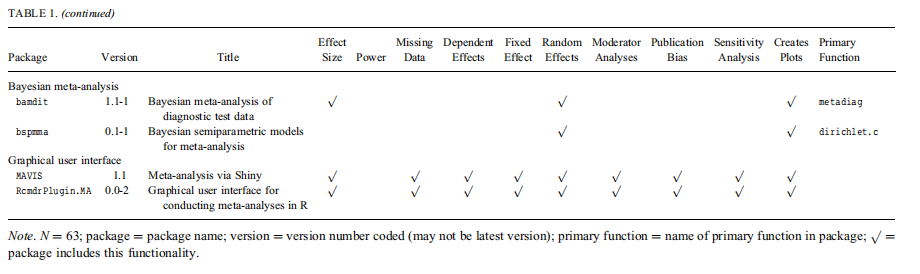
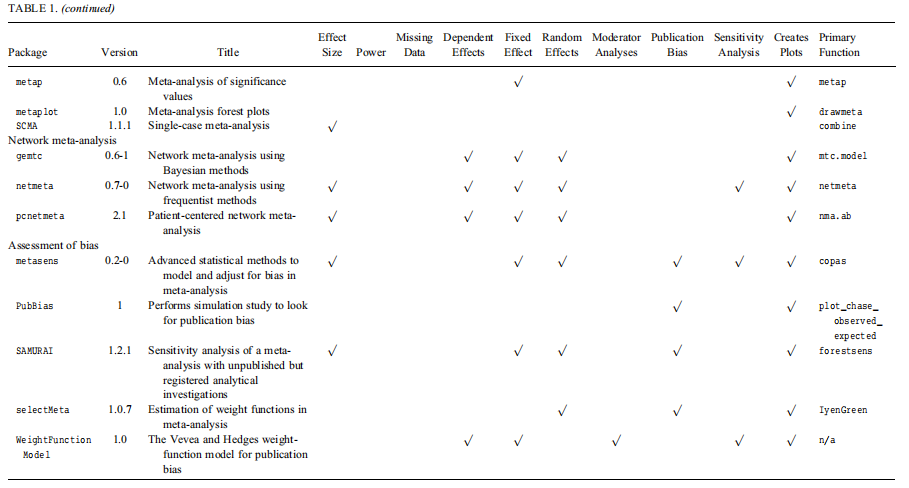
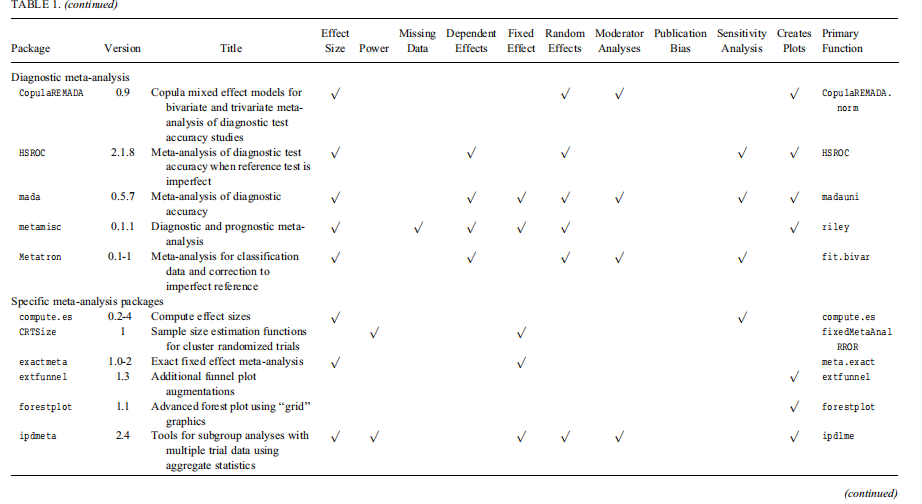
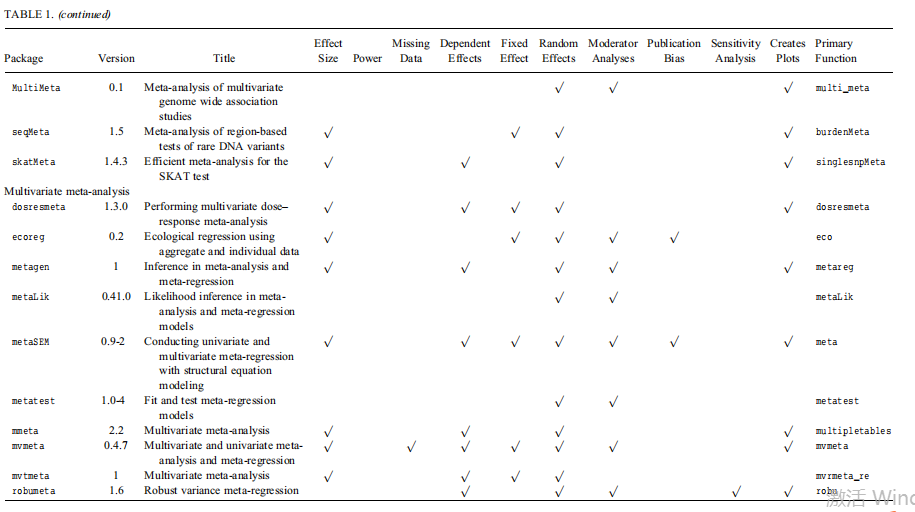
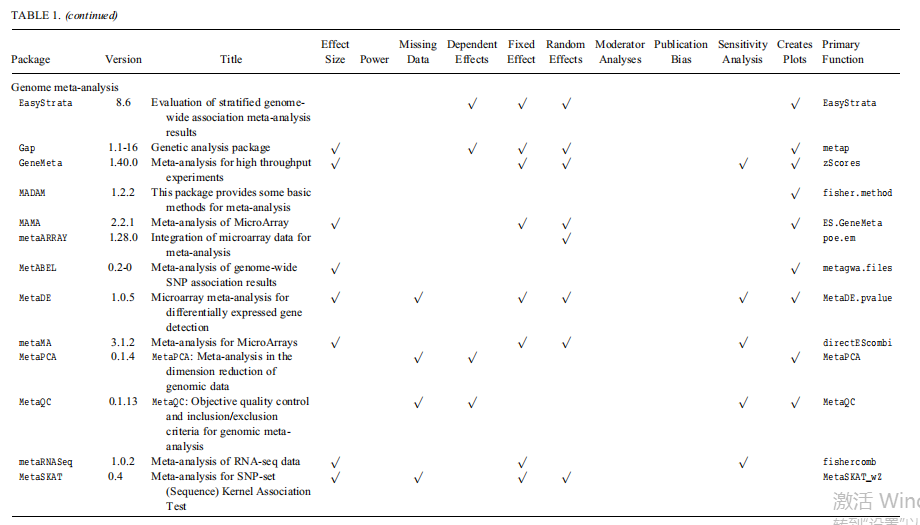
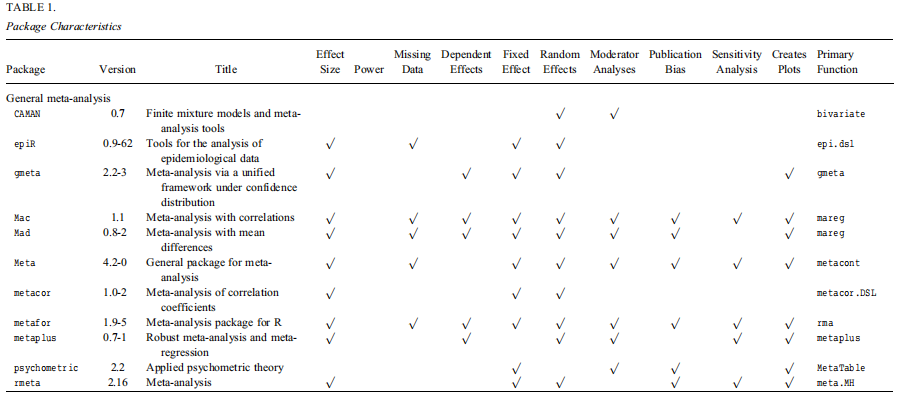


图1 R程序包创建的年份分布

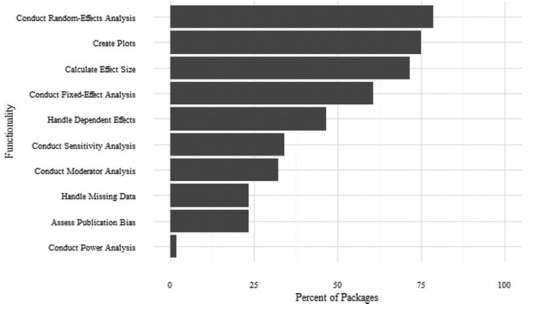


图2所有R元分析程序包的功能

这些不同的程序包在使用中也有很大差异，2016年3月元分析程序包最高下载量是epiR(Stevenson,2015)，平均日下载量高达112.4次（SD=26.06）。根据所有数据显示，平均每个程序包下载量为8.56次（SD=11.19）。然而与其他常用的R程序包相比，元分析程序包的下载量则相对较少。比如，2016年3月使用最普遍的绘图元分析包ggplt22 (Wickham, 2009)日下载量为8467.13次。

## 4. 元分析R程序包的分类描述

### 4.1 General Meta-analytic Packages

在63个程序包中有11个程序包为General Meta-analytic 程序包。此组程序包与传统元分析的大部分程序相同，只适用于一般性问题，不适用于某一特定领域的问题。因此，这11 个程序包更适合用于解决基础性元分析问题。例如，这些程序包只用几行简单的代码就可以计算固定效应或随机效应的加权平均值。

Metafor程序包是其中最受欢迎的程序包(Viechtbauer, 2015)。Metafor有许多功能，可以解决使用者的大部分问题，从计算效应量到估计随机效应模型再到制作综合结果图。另外，它还可以使用listwise deletion解决缺失值问题，使用多层建模处理层级相关的效应值，也可以进行调节效应分析、敏感性分析以及评价出版偏倚等问题。

除Metafor之外，还有几个其他几个程序包具有较为全面的功能。MAc(Del Re & Hoyt, 2015a) 和MAd(Del Re &Hoyt, 2015b)程序包分别用相关和标准平均差的效应值估计主效应的大小。因此，他们具有Metafor没有的一些功能，例如对效应值进行校正。类似的，metacorr (Laliberte, 2015)主要关注的也是相关效应值，但它不能处理层级相关的效应值。此外，与MAc相比，metacorr 的资料也相对较少。Meta(Schwarzer, 2015)和rmetaa (Lumley, 2012) 程序包也可以提供一些全面的的元分析功能，但并不能处理层级相关的效应值。Yang, Shi, and Xie (2016)的gmeta程序包包含多元元分析功能，可以处理层级相关的效应值。而产生于1999年的rmeta程序包是最古老的一种R程序包 (R Core Team, 2015)。

少数几个一般性的程序包包含了多种功能，但它们只用于某一特定领域。比如说，epiR程序包（Stevenson，2015）用于流行病学元分析研究，psychometric程序包（Fletcher，2015）用于应用心理测量学理论的元分析研究，例如，效度概化以及生成元分析模型的函数；metaplus程序除了其他通用功能外，还可以进行稳健元回归研究（Beath，2015）。

### 4.2 Genome Meta-Analytic Packages

基因和微阵列（microarray）元分析领域包含的程序包数量最多（n=16）。这些程序包都只针对特定领域，不适用于一般化的元分析，比如教育领域和社科领域。

其中5个程序包主要用于微阵列数据元分析，分别是MADAM (Kugler, Mueller, Ecker, & Graber, 2015)，MAMA (Ihnatova, 2015)，metaARRAY (Ghosh & Choi, 2015)，MetaDE (Wang, Li, & Tseng, 2015)和MetaMA (Marot, 2015)。MAMA程序包可以计算效应量，计算固定效应模型或随机效应模型，还可以制作一些数据图，但此数据包使用时需要依靠GeneMeta、metaMA、metaARRAY这三个程序包。metaARRAY数据包主要用于对贝叶斯框架中大规模的微阵列数据进行元分析，并且包括使用公共可用数据对程序包功能的编写进行演示 (Ghosh & Choi, 2015)。MetaDE程序包为使用者提供了12种关于基因表达差异性（gene express）的元分析选项，使用者可以根据结果变量选择检验统计量，然后进行修正。MetaDE程序包可能是唯一一个可以处理缺失值的微阵列程序包，当缺失值少于30%时，它将自动填充缺失数据；当缺失数据多于30%时，它将自动忽略这些序列。MetaMA程序包将结合配对数据和非配对数据中的p值或t检验统计量来寻找基因表达差异性（Marot & Bruyere, 2015）。这里有一个使用公共数据对MetaMA程序包进行的演示（(Marot & Bruyère, 2015）。Marot，Jaffrezic, and Rau的metaRNAseq程序包也有相似的功能，主要用来进行RNA序列实验（sequencing experiments）。

此类别中的其他程序包也执行不同的特定任务。Kangand Tseng（Kang，Tseng，2015a）的MetaPCA程序包进行因子降维程序；他们的另一个项目metaQC（Kang，Tseng，2015b）程序包主要力求改进数据的纳入和排除标准等。Lee (2015)的 metaSKAT主要用于“对稀有变异体基于基因的元分析”(p. 1)。Voorman, Brody, andLumley (2015)的skatMETA 和Voorman, Brody, Chen, and Lumley (2015)的seqMETA有相似的功能。metABEL (Struchalin & Aulchenko,

2015)程序包主要用于对“定量或二元试验与SNP之间的全基因组关联扫描 ”进行元分析(p. 1)。MultiMETA (Vuckovic, 2015) 和 EasyStrata (Winker, 2015)专门处理全基因组数据。Zhao (2015)的 gap程序包主要对通用基因组测序进行元分析。

### 4.3 Multivariate Meta-Analytic Packages

Multivariate Meta-Analytic Packages程序包主要是多变量元分析类型，对应用型和高级研究者可能会有帮助。该类型的程序包(n=10)共享允许使用多个协变量或结果的多变量统计过程。

这类程序包进一步可以划分为能够处理相依型效应值（dependent effect size）的和不能处理相依型效应值的两类。在Multivariate Meta-Analytic Packages中，有7个程序包可以处理相依型效应值，其中3个主要进行一般的和传统的多变量元分析，mmeta（Luo，Chen，Su，& Chu，2015）、mvmeta（Gasparrini，2015）和mvmeta（Chen，2015）用于分析已知或可以估计层级相关效应值之间的协变量。例如，mvmeta程序包就是使用层级相关效应值中的协变量来调整结果效应量、结合多时间点拟合模型以及建立网络元分析模型或进行诊断测试。

另外3个程序包dosremata (Crippa, 2015)、metaSEM(Cheung, 2015)和robumeta (Fisher

& Tipton, 2015)用来计算效应量的相依性，它们并不要求计算效应量中的协变量，而是计算标准误的无偏估计。dosremata程序包主要用 Greenland and Longnecker (1992)和 Hamling and colleagues (2008)提出的方法估计协方差变异。robumeta程序包使用 Hedges, Tipton, and Johnson (2010)提出的稳健方差估计技术估计标准误。metaSEM使用结构方程模型的框架解释研究内的相依性，在metaSEM程序包的开发者网站上就可以获得一般性的说明和示例。这些程序包的好处是在不知道精确的效应值之间的协变量情况下，可以同时综合多变量相关效应值。尽管这些程序包都具有绘图功能，但robumeta (Fisher & Tipton, 2015) 可以根据研究将效应值进行分组并产生传统的森林图，这是这个程序包的独特之处。

第二类程序包主要用来建立元回归模型，但不能处理相依型效应值（n=3），因此，在使用metagen、metaLik、metatest这三个程序包时，分析者必须从每个研究中选择一个效应量纳入到模型中。metagen程序包具有执行几种不同类型的基于推理的元回归模型的能力。metaLik程序包可以“使用有符号对数似然比测试（the signed log-likelihood ratio test）对固定效应向量的标量分量的假设检验''(p. 9)，还具有仿真功能。metatest程序包为元回归模型中的各种参数产生一系列的测试统计信息。

### 4.4 Diagnostic Test Accuracy Meta-Analytic Packages

虽然其他几个程序包的作者有具体说明可以使用他们的程序包来完成诊断性测验的精度元分析（例如metafor、mvmeta），但有5个专门为诊断元分析而设计的程序包，其中3个程序包只能完成一般的诊断元分析，另外两个程序包可以解决专门的数据问题。Mada（Dobler，2015）程序包可以估计固定和随机效应分析、进行调节效应分析和敏感性分析以及制作元分析结果图和接收者工作特征（ROC）曲线图。Metamisc（Debray，2015）数据包也可以处理预后测验（prognostic testing），它能采用频率或叶贝斯方法处理研究内的相依关系，但它不能进行调节标量分析和敏感性分析。CopulaREMADA（Nikoloulopoulos，2015）程序包主要使用的是二元和三元变量诊断测验的精确元分析。HSROC (Schiller & Dendukuri, 2015) 和Metatron (Huang, 2015)主要在参考测验不理想的情况下进行的诊断测验精度元分析。Metatron (Huang, 2015)不仅可以进行传统的诊断元分析，也可以估计特点研究对平均点估计的影响。此外，HSROC程序包可以使用分层建模技术同时分析研究内和研究间的变异，因此，这个程序包对由于分层聚类而导致的标准误非常有用。

### 4.5 Network Meta-Analysis Packages

网络元分析程序包主要有三个：gemtc（van Valkenhoef & Kuiper，2014）、pcnetmeta（Lin，Zhang，2015）和netmeta（Krahn,2015)。其中，gemtc程序包使用一般性的贝叶斯方法，而pcnetmeta程序包使用特殊性的贝叶斯方法进行以患者为中心的网络元分析，netmeta程序包是在图论的基础上使用频率论方法进行网络元分析。这三个程序包全部都有制图功能，包括森林图、网络图（network graph）和净热图（net heat plot）。pcnetmeta程序包依赖于另一个Gibbs 采样器进行参数估计，且这个采样器需要在网上单独下载（http://mcmc-jags.sourceforge.net）。有关网络元分析程序包的更多的讨论以及该功能的教程和解释，请参阅Neupane等人（2014）关于这些程序包的报告。

### 4.6 Bayesian Meta-Analysis Packages

有两个程序包专门针对采用贝叶斯方法的元分析。Bamdit（Verde & Sykosch，2015）程序包能够对诊断测验的数据进行随机效应建模，计算效应值，并绘制贝叶斯ROC曲线图。它也可以通过ggplot2（Wickham，2009）在元分析中绘制每个研究的真阳性率和假阳性率图，并能使用不同的颜色来显示不同的亚组。第二个使用贝叶斯方法的是bspmma程序包，此程序包可以估计非参数和半参数随机效应元分析，但其制图能力有限。

### 4.7 Assessment of Bias Packages

有5个程序包用来评价出版偏倚，其中4个用来评价元分析中出版偏倚的可能性。这4个程序包分别是PubBias（Thornley，2015）、SAMURAI（Noory 2015）、selectMeta（Rufibach，2015）和WeightFunctionModel（Vevea & Coburn，2015）。PubBias程序包用于评估被观测的积极效应的数量是否大于费舍Z检验的期望值。SAMURAI程序包可以补充分类变量和连续变量的缺失值，并绘制森林图和漏斗图。selectMeta程序包通过使用非参数和参数权重函数对潜在偏差进行建模，从而解决出版偏倚问题。WeightFunctionMode程序包可以使用网络用户界面（https：//vevealab.shinyapps.io/WeightFunctionModel/）来运行R程序包，在这个程序中，用户需要完成一系列数据问题，选择变量（必要的话包括调节变量），建立期望的p值临界点，然后计算出Vevea和Hedges（1995，1996）出版偏倚的权重函数模型，并产生一个随机漏斗图。

这类中的最后一个程序包——metasens（Schwarzer 2015）是评估潜在偏倚来源的一个扩展包，它可以进行有限元分析，也可以通过漏斗图、等高线图、实验组效应以及残差选择偏倚的p值图来评价潜在的偏倚。

### 4.8 Packages With Specific Functionality

有9个程序包是对特殊数据类型或进行元分析某一特殊方面。Compute.es (Del Re,2015a)主要用来计算6种受普遍的效应量（Cohen’s d, Hedges’g, r, Fisher’s z, odds ratio, log odds ratio）以及效应量的变异、置信区间和p值。R中的MAc程序包和MAVIS GUI中就采用了这种个程序包。metap (Dewey, 2015)程序包可以通过不同的方法结合P值，如logit、平均P值、logs的和、p值和、z值和投票计数以及Wilkinson法。

有3个程序包扩展了现有的制图功能：

1. 扩展漏斗图（Langan, sutton, Higgins,&Gegory,2015），增加了统计显著性和异质性等高图；
2. 森林图（Gordon & Lumley, 2015），使用网格图形获得更高级的森林图选项；
3. 元图，使用R程序包中的meta, metafor和rmeta（Murrell, Scott, & Lu, 2014）输出。

还有三个程序包用于特殊数据类型的元分析，exactmeta (Yu & Tian, 2015)程序包对稀有事件数据进行固定效应元分析，ipdmeta(Kovalchik,2015)程序包对独立患者数据进行元分析，SCMA(Bulte, 2015)程序包对个案数据进行元分析。

最后，CRTsizee (Rotondi, 2015)程序包包含了元分析功效的估计。

### 4.9 GUI Applications

有两个程序包具有GUI 功能：MAVIS（Hamilton，Mizumoto, Aydin,&Coburn,2015）和RcmdrPlugin。GUI程序的优点在于使用者不必知道如何编码或其他额外的知识。MAVIS程序包和SPSS有点类似，都是需用使用者点击一系列下拉菜单来完成。这个程序包可以预测固定效应模型和随机效应模型，还可以进行基础的调节效应分析和制图。RcmdrPlugin.MA程序包是Rcmdr程序包的补充。使用者首先要下载Rcmdr GUI，然后再添加元分析程序包。进入用户界面后从菜单中选择元分析选项。那些不想学习R程序语言的人们可以通过简单的点击选项来使用高水平的元分析功能。

## 5. 两种元分析程序包的使用指南

接下来我们提供了两个程序包的简要指南：metafor（Viechtbauer, 2015）程序包和robumeta（Fisher & Tipton, 2015）程序包。虽然程序包的帮助文件中提供了很多例子和功能说明，在Stack Overflow等站点上可以发现活动用户的不断增长(Stack Exchange,

2016)。但根据我们的经验来看，对于初学者来说，学习一个新的R程序包仍旧有些困难。因此这部分的目的就是向R用户和元分析研究的新手提供足够的细节，以便可以进行基本的元分析过程。另外，我们还介绍了与传统元分析模型相比，稳健方差估计（robust variance estimation）的优点。稳健方差估计允许元分析人员在一个模型中合成所有可用的效果大小，即使在特定研究中有多个可用的效果大小。目前为止，针对这种情况，大多数研究者都是选择以下方法的其中一种：（1）每篇研究只选择一个效应值进行综合分析；（2）在进行综合分析之前，对每个研究内的多个效应值进行平均；（3）将元分析分解为多个较小的、特定结果的元分析。如果使用稳健方差估计就不会面临这些问题。

接下来我们以最近的一篇关于减少中学生青少年约会暴力的预防方案的系统回顾和元分析的文章为例进行说明（De La Rue, Polanin, Espelage, & Pigott, 2016）。该研究对感兴趣的是青少年约会暴力知识的结果，并将有效的治疗效果的效应值编码为负向效应量（负性效应量表示约会暴力率的下降）。数据集包括13个独立样本和25个效应值。表2提供了所有运行R需要的编码信息。除了效应值外，示例数据还包括效应值方差（var），研究名称（Study.ID），effect size ID（ES.ID），作者（author）以及随访（Follow.Up）。同时，列出了三个调节变量，分别为：出版时间（DOP）、研究是否随机分配（Random）以及干预组的男性百分比（Perc.Males）。结果仅供演示，对完整结果感兴趣的读者可以查阅原文。我们在每一部分前做了代码指导，在编码中使用#来表示注释信息。

表2青少年约会暴力元分析数据集

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Study.ID | ES.ID | DOP | *g* | var | Follow.Up | Random | Perc.Males1 |
|  |  |  |  |  |  |  |  |
| 1 | 2 | 1998 | 0.01 | 0.015 | 1 | 1 | 48.9 |
| 2 | 3 | 2006 | 0.707 | 0.002 | 0 | 1 | 48.8 |
| 2 | 4 | 2006 | 0.349 | 0.002 | 1 | 1 | 48.8 |
| 3 | 5 | 1991 | 0.279 | 0.007 | 0 | 0 | NA |
| 3 | 6 | 1991 | 0.211 | 0.007 | 0 | 0 | NA |
| 4 | 7 | 1996 | 0.286 | 0.009 | 0 | 0 | NA |
| 4 | 8 | 1996 | 0.003 | 0.009 | 1 | 0 | NA |
| 5 | 9 | 1997 | 0.272 | 0.009 | 0 | 1 | 49.3 |
| 6 | 10 | 2002 | 0.021 | 0.007 | 0 | 0 | 45.03 |
| 6 | 11 | 2002 | 0.047 | 0.007 | 1 | 0 | 45.03 |
| 6 | 12 | 2002 | 0.348 | 0.007 | 0 | 0 | 45.03 |
| 6 | 13 | 2002 | 0.072 | 0.007 | 1 | 0 | 45.03 |
| 6 | 14 | 2002 | 0.355 | 0.012 | 0 | 0 | 45.91 |
| 6 | 15 | 2002 | 0.036 | 0.012 | 1 | 0 | 45.91 |
| 6 | 16 | 2002 | 0.389 | 0.012 | 0 | 0 | 45.91 |
| 6 | 17 | 2002 | -1.739 | 0.016 | 1 | 0 | 45.91 |
| 7 | 18 | 2000 | -0.121 | 0.001 | 0 | 0 | 50.5 |
| 8 | 19 | 2000 | -0.492 | 0.014 | 0 | 1 | 42.8 |
| 9 | 20 | 2010 | -0.206 | 0.012 | 0 | 1 | 52 |
| 9 | 21 | 2010 | -0.194 | 0.01 | 1 | 1 | 52 |
| 10 | 22 | 2012 | -0.165 | 0.011 | 0 | 1 | 54.2 |
| 10 | 23 | 2012 | -0.119 | 0.006 | 1 | 1 | 54.2 |
| 11 | 24 | 2001 | -1.233 | 0.199 | 0 | 0 | 45.8 |
| 11 | 25 | 2001 | -1.425 | 0.209 | 1 | 0 | 45.8 |

*Note*. DOP ¼ date of publication.

### 5.1程序包的安装和加载

首先开始清理R环境并加载元分析程序包。如果是第一次使用metafor或robumeta程序包，需要安装程序包。在R程序主页面输入：install.packages(“metafor”)，然后开始启动安装，最后一段代码（sessionInfo）列出了接下来可以运行的R和程序包的版本，如果代码无法运行，可以换一个版本。

rm (list = ls())

}]

library (metafor)

## Loading required package: Matrix

## Loading ‘metafor’ package (version 1.9-8). For an overview

## and introduction to the package please type: help(metafor).

library (robumeta)

## Loading required package: grid

sessionInfo ()

## R version 3.2.2 (2015-08-14)

## Platform: x86\_64-w64-mingw32/x64 (64-bit)

## Running under: Windows 7 x64 (build 7601) Service Pack 1

##

## locale:

## [1] LC\_COLLATE = English\_United States.1252

## [2] LC\_CTYPE = English\_United States.1252

## [3] LC\_MONETARY = English\_United States.1252

## [4] LC\_NUMERIC = C

## [5] LC\_TIME = English\_United States.1252

##

## attached base packages:

## [1] grid stats graphics grDevices utils datasets methods

## [8] base

##

## other attached packages:

## [1] robumeta\_1.6 metafor\_1.9-8 Matrix\_1.2-2

##

## loaded via a namespace (and not attached):

## [1] magrittr\_1.5 tools\_3.2.2 htmltools\_0.2.6 yaml\_2.1.13

## [5] stringi\_1.0-1 rmarkdown\_0.8.1 knitr\_1.11 stringr\_1.0.0

## [9] digest\_0.6.8 lattice\_0.20-33 evaluate\_0.8

### 5.2 上传和准备数据

接下来，使用”read.csv”功能上传数据集（date set）（date\_TDV.csv）。使用者可以通过”head”和”str”功能显示数据集。确定数据集上传完成后，我们使用”Follow.Up”变量将数据集划分为两个部分。编码为0的是后测结果, 而1行是随访结果。通常我们不会同时对后测和随访结果同时进行综合分析，而是分开计算。

这部分我们详细介绍了“传统”元分析（所有效应值都是独立变量）所需的步骤。在5.5部分，说明了使用稳健方差估计的最新元分析技术所必需的步骤。我们建议读者，尽量使用最新的元分析技术。

dat <- read.csv ("Data\_TDV.csv", header = T)

head (dat)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ## | Study.ID | | ES.ID | | Author DOP | | Follow.Up | g | va r |
| ## 1 | 105 | | 1 | | Foshee, V. A. 1998 | | 0 | 0.0100000 | 0.014621856 |
| ## 2 | 105 | | 2 | | Foshee, V. A. 1998 | | 1 | 0.0100000 | 0.014621856 |
| ## 3 | 110 | | 3 | | Jaycox, L.H. 2006 | | 0 | -0.7074087 | 0.001864498 |
| ## 4 | 110 | | 4 | | Jaycox, L.H. 2006 | | 1 | -0.3487226 | 0.001781894 |
| ## 5 | 111 | | 5 | | Jones, L.E. (1) 1991 | | 0 | -0.2792240 | 0.007212470 |
| ## 6 | 112 | | 6 | | Jones, L.E. (2) 1991 | | 0 | -0.2107261 | 0.007387867 |
| ## Random Perc.Males | | | | | | | | | |
| ## 1 | 1 | | 48.9 | | | | | | |
| ## 2 | | 1 | | 48.9 | |
| ## 3 | | 1 | | 48.8 | |
| ## 4 | | 1 | | 48.8 | |
| ## 5 | | 0 | | NA | |
| ## 6 | | 0 | | NA | |
|  | |  | |  | |

str (dat)

## ‘data.frame’: 25 obs. of 9 variables:

## $ Study.ID : int 105 105 110 110 111 112 113 113 114 117 …

## $ ES.ID : int 1 2 3 4 5 6 7 8 9 10 …

## $ Author : Factor w/ 12 levels ‘Foshee, V. A.",..: 1 1 2 2 3 4 5 5 6 7 …

## $ DOP : int 1998 1998 2006 2006 1991 1991 1996 1996 1997 2002 …

## $ Follow.Up : int 0 1 0 1 0 0 0 1 0 0 …

## $ g : num 0.01 0.01 -0.707 -0.349 -0.279 ...

## $ var : num 0.01462 0.01462 0.00186 0.00178 0.00721 ...

## $ Random : int 1 1 1 1 0 0 0 0 1 0 ...

## $ Perc.Males : num 48.9 48.9 48.8 48.8 NA ...

dat.0 <- subset(dat, Follow.Up == "0")

dat.1 <- subset(dat, Follow.Up == "1")

### 5.3 准备传统元分析所需的数据集

进行传统元分析，首先要确保所有的效应值都是独立变量。当单个研究中对同一样本用多个效应值表示时，确保效应值独立性的一个办法就是对效应值及其方差进行平均。我们对具多个效应值的研究进行再分组（4-5列和6-7列），其次，对每个研究内的多个效应值和方差进行平均，并创建一个新的变量来表示（如“avg.g.117”）。最后，用这些平均值替换统计信息列表中的原始效应值和方差。

最后，将具有新的效应值和方差的行与原始数据集结合(rbind(dat.1, dat.1.117, dat.1.118))，并将没有用到的原始数据列删除(dat.1[-c(4,5,6,7,12,14),])，剩下的就是具有独立效应值和方差的元分析数据集。

#create a small dataset with the studies that include multiple effect sizes

dat.1.117 <- dat.1[c(4,5),]

dat.1.118 <- dat.1[c(6,7),]

#average the effects within each study

avg.g.117 <- mean(dat.1.117$g)

avg.var.117 <- mean(dat.1.117$var)

avg.g.118 <- mean(dat.1.118$g)

avg.var.118 <- mean(dat.1.118$var)

#replaces the multiple effect sizes with the average effect size

dat.1.117[1,6] <- avg.g.117

dat.1.117[1,7] <- avg.var.117

dat.1.118[1,6] <- avg.g.118

dat.1.118[1,7] <- avg.var.118

dat.1 <- rbind(dat.1, dat.1.117, dat.1.118)

dat.1 <- dat.1[-c(4,5,6,7,12,14),]

### 5.4 使用metafor进行传统元分析

数据集完成后，开始进行传统元分析。下面进行了四个不同的分析，第一个是没有调节变量的简单元分析，其中”yi”表示效应量，”vi”表示方差。”date”语句告诉R运行哪个数据集，“method”告诉R选择运行哪个元分析模型。在这里，我们使用最大似然比（ML）估计随机效应模型。编码完成后，我们点击”summary(run1)”显示输出。

输出结果中列出了所运行的元分析所有的相关信息。在”Model Results”标题下，我们可以看到平均效应量（估计）、平均效应量的标准误（se）、z检验的显著性（pval）、95%置信区间的上下限（“ci.lb”,”ci.ub”）。在”Test for Heterogeneity”下可以看到模型的异质性信息。在此例子中，结果显示平均效应值具有显著的统计学意义，而异质性结果不显著。

run.1 <- rma.uni (yi = g, #effect size

vi = var, #variance

data = dat.1, #dataset

method = "ML" #model estimation )

summary(run.1)

##

## Random-Effects Model (k = 8; tau^2 estimator: ML)

##

## logLik deviance AIC BIC AICc

## -3.8449 28.7496 11.6897 11.8486 14.0897

##

## tau^2 (estimated amount of total heterogeneity): 0.0874 (SE = 0.0510)

## tau (square root of estimated tau^2 value): 0.2957

## I^2 (total heterogeneity / total variability): 91.84%

## H^2 (total variability / sampling variability): 12.26

##

## Test for Heterogeneity:

## Q(df = 7) = 62.4666, p-val < .0001

##

## Model Results:

##

## estimate se zval pval ci.lb ci.ub

## -0.2644 0.1146 -2.3072 0.0210 -0.4890 -0.0398 \*

##

## ---

## Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘‘ 1

进行调节分析时，有多种选择。第二、第三、第四种模型分别进行不同类型的调节效应分析。第二个模型使用”mods”进行单因素方差分析的调节分析。末尾的“-1”表示R删除的截距，系数代表模型中每个水平下的平均效应量。”Test of Moderators”提供了Q检验(Q-between statistic)的信息，结果显示组间差异显著（Q-between=6.29，p=.04）。

run.2 <- rma.uni (yi = g,

vi = var,

data = dat.1,

method = "ML",

mods =~factor(Random) — 1 #moderator argument)

summary (run.2)

##

## Mixed-Effects Model (k = 8; tau^2 estimator: ML)

##

## logLik deviance AIC BIC AICc

## -3.4152 27.8903 12.8304 13.0687 18.8304

##

## tau^2 (estimated amount of residual heterogeneity): 0.0847 (SE = 0.0495)

## tau (square root of estimated tau^2 value): 0.2910

## I^2 (residual heterogeneity / unaccounted variability): 90.97%

## H^2 (unaccounted variability / sampling variability): 11.07

##

## Test for Residual Heterogeneity:

## QE(df = 6) = 61.7007, p-val < .0001

##

## Test of Moderators (coefficient(s) 1, 2) :

## QM(df = 2) = 6.2936, p-val = 0.0430

##

## Model Results:

##

## estimate se zval pval ci.lb ci.ub

## factor(Random)0 -0.3797 0.1688 -2.2495 0.0245 -0.7106 -0.0489 \*

## factor(Random)1 -0.1689 0.1521 -1.1105 0.2668 -0.4669 0.1292

##

## ---

## Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

第三、第四个模型是亚组分析。在本示例中，数据集被划分为随机组和非随机组。我们再次使用”summary”功能查看结果。结果显示，非随机分配组的平均效应量相对较大，但并不显著（p=0.09）；虽然随机分配组平均效应量相对较小，但在统计学意义上具有显著性（p=0.01）。值得注意的是，这些元分析平均效应量与模型二不同，因为模型二包括研究内的方差，而模型三和模型四不包括 (关于调节因子和亚组分析的更多信息可以参考 Borenstein, Hedges, Higgins, & Rothstein, 2010)。

run.3 <- rma.uni (yi = g,

vi = var,

data = dat.1,

method = "ML",

subset = (Random == "0") #subgroup argument)

run.4 <- rma.uni (yi = g,

vi = var,

data = dat.1,

method = "ML",

subset = (Random == "1") #subgroup argument)

summary (run.3)

##

## Random-Effects Model (k = 4; tau^2 estimator: ML)

##

## logLik deviance AIC BIC AICc

## -3.4159 14.9449 10.8317 9.6043 22.8317

##

## tau^2 (estimated amount of total heterogeneity): 0.2358 (SE = 0.1913)

## tau (square root of estimated tau^2 value): 0.4856

## I^2 (total heterogeneity / total variability): 94.35%

## H^2 (total variability / sampling variability): 17.70

##

## Test for Heterogeneity:

## Q(df = 3) = 48.4943, p-val < .0001

##

## Model Results:

##

## estimate se zval pval ci.lb ci.ub

## -0.4533 0.2631 -1.7231 0.0849 -0.9690 0.0623

##

## ---

## Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

summary(run.4)

##

## Random-Effects Model (k = 4; tau^2 estimator: ML)

##

## logLik deviance AIC BIC AICc

## 2.3953 8.1562 -0.7906 -2.0180 11.2094

##

## tau^2 (estimated amount of total heterogeneity): 0.0114 (SE=0.0125)

## tau (square root of estimated tau^2 value): 0.1069

## I^2 (total heterogeneity / total variability): 65.63%

## H^2 (total variability / sampling variability): 2.91

##

## Test for Heterogeneity:

## Q(df = 3) = 13.2064, p-val = 0.0042

##

## Model Results:

##

## estimate se zval pval ci.lb ci.ub

## -0.1921 0.0676 -2.8395 0.0045 -0.3246 -0.0595 \*\*

##

## ---

## Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ‘ 1

### 5.5 使用robumeta程序包进行稳健方差估计

接下来，我们使用”dat.0”数据集（仅有后测效应值）进行稳健方差元分析。”dat.0”数据集是用5.2部分的“subset”功能创建的。因为我们使用的是稳健方差估计模型，不需要确定数据是不是独立效应值，因此，我们不需要像5.3部分那样操作数据集。该模型假设研究内各效应值之间的相依性为常数。我们可以通过敏感性分析确定这个相关系数对结果不会产生显著的影响。

首先，我们先运行一个不包含调节变量的元分析。针对“robumeta”程序包，我们必须告诉R程序计算效应值的公式什么。这个公式在之后的调节变量分析中也会用到。如示例中所示，公式简单地表明哪一列市效应值列，并且在第1列上进行回归分析，因此语句为  
“formula =g~1”。效果值方差用“var.eff.size”语句表达，“data”显示数据所在的程序包，”studynum”告诉软件哪一列拥有研究水平的ID（which column has the study-level IDs）。每个研究肯会有多个效应值，但来自同一个研究的效应量应该有相同的”Study.ID”。最后一个语句modelweights=’CORR’表示由于每个研究有多个效应值导致的相关。另外，“modelweights”中的''HIER''选项主要用于效应值来源于某个分层集群（例如，作者团队、实验室或国家），不适用于此示例。最后，我们使用”print”语句来查看平均效应量是否达到了统计性显著。

我们可以运行”sensitivity(run.5)”语句，来检验研究内相关对结果是否敏感。 这个函数的结果范围在0到1之间，输出结果显示了针对每个研究内相关性变化的模型。在当前例子中，p值的变化并没有改变平均值（g=-.2159），所以我们可以认为研究内的相关并不影响结果。

最后6、7两个模型进行了调节分析。模型6检验了随机组和非随机组之间的差异，结果显示两组之间的效应量并没有显著差异（p=.23）。”Note”中的“如果*df<4,*则结果不可靠”是指回归系数相关的自由度。Tanner-Smithet al. (2016)表示，自由度越小，结果越不稳定，因此需要谨慎。

模型7是多元回归模型，包括随机二元变量和”Perc.Males”连续变量。结果表明，此调节因子不是影响效应值大小显著预测因子。

run.5 <- robu (formula = g \* 1, #effect size

var.eff.size = var, #variance

data = dat.0, #dataset

studynum = Study.ID, #study-level ID

modelweights = "CORR" #RVE model estimation)

print (run.5)

## RVE: Correlated Effects Model with Small-Sample Corrections

##

## Model: g ~ 1

##

## Number of studies = 13

## Number of outcomes = 15 (min = 1 , mean = 1.15 , median = 1 , max = 2)

## Rho = 0.8

## I.sq = 94.15285

## Tau.sq = 0.09484842

##

## Estimate StdErr t-value dfs P(|t|>) 95% CI.L 95% CI.U Sig

##1 intercept -0.216 0.085 -2.54 11.6 0.0265 -0.402 -0.03 \*\*

##---

## Signif. codes: < .01 \*\*\* < .05 \*\* < .10 \*

## ---

## Note: If df < 4, do not trust the results

sensitivity(run.5)

## RVE: Correlated Effects Model with Small-Sample Corrections

## Model: g ~ 1

##

## Sensitivity Analysis

##

## Rho = 0 Rho = 0.2 Rho = 0.4 Rho = 0.6 Rho = 0.8

## intercept Coefficient -0.2159 -0.2159 -0.2159 -0.2159 -0.2159

## Std. Error 0.0850 0.0850 0.0850 0.0850 0.0850

## Tau.sq Estimate 0.0948 0.0948 0.0948 0.0948 0.0948

## Rho = 1

## -0.2159

## 0.0850

## 0.0949

run.6 <- robu (formula = g ~factor(Random) #moderator argument,

var.eff.size = var,

data = dat.0,

studynum = Study.ID,

modelweights = "CORR")

print (run.6)

## RVE: Correlated Effects Model with Small-Sample Corrections

##

## Model: g ~ factor(Random)

##

## Number of studies = 13

## Number of outcomes = 15 (min = 1 , mean = 1.15 , median = 1 , max = 2)

## Rho = 0.8

## I.sq = 90.30503

## Tau.sq = 0.06545406

##

## Estimate StdErr t-value dfs P(|t|>) 95% CI.L 95% CI.U

## 1 intercept -0.116 0.114 -1.01 5.47 0.354 -0.401 0.17

## 2 factor.Random.1 -0.200 0.158 -1.27 10.36 0.232 -0.550 0.15

## Sig

## 1

## 2

## ---

## Signif. codes: < .01 \*\*\* < .05 \*\* < .10 \*

## ---

## Note: If df < 4, do not trust the results

run.7 <- robu (formula = g ~ factor(Random) + scale(Perc.Males),

var.eff.size = var,

data = dat.0,

studynum = Study.ID,

modelweights = "CORR")

print (run.7)

## RVE: Correlated Effects Model with Small-Sample Corrections

##

## Model: g ~ factor(Random) + scale (Perc.Males)

##

## Number of studies = 10

## Number of outcomes = 12 (min = 1, mean = 1.2, median = 1, max = 2)

## Rho = 0.8

## I.sq = 93.06073

## Tau.sq = 0.113052

##

## Estimate StdErr t-value dfs P(|t|>) 95% CI.L 95% CI.U

## 1 intercept -0.00981 0.220 -0.0446 2.67 0.968 -0.763 0.743

##2 factor.Random.1 -0.32827 0.236 -1.3890 3.89 0.239 -0.992 0.335

##3 scale.Perc.Males. 0.05898 0.056 1.0533 2.52 0.383 -0.140 0.258

## Sig

##1

##2

## 3

## ---

## Signif. codes: < .01 \*\*\* < .05 \*\* < .10 \* ## ---

## Note: If df < 4, do not trust the results

## 6. 讨论

对于研究者和统计学家来说，R统计环境是一个非常普遍而且有用的工具。在过去的十年里，R程序包变得越来越受欢迎，用户创建以及R开放源代码的增加也促使R的功能和效用也不断增多。元分析社区通过添加许多特定于元分析或包括元分析功能的附加包对此增长也作出了贡献。本文章的目的是搜索到全部的元分析R程序包，然后描述它们的功能和用法，最后提供两个热门的程序包的使用指南。

对于初级学者来说，我们推荐使用一般性的元分析程序包，其中，metafor(Viechtbauer, 2015)程序包具有很多功能，是元分析R程序包中下载量最高的且灵活性最高的程序包，很适合对元分析感兴趣的研究者使用。还有许多其他具有特殊功能的程序包也可以下载，包括用于诊断测试精度元分析的程序包、用于基因数据类型的程序包以及贝叶斯元分析程序包等。

### 6.1 局限性

当前研究的局限之一是我们只对进行元分析的统计软件进行综述，而没有对可能用于进行系统评价和提取元分析数据的软件进行综述。元分析是一个系统搜索文献的结果，这一过程包含了大量的数据管理。R程序包在数据管理过程中是一个有用的工具，但目前我们还没有任何R程序包可以彻底解决这个问题。。在理想情况下，程序包对管理系统评价和元分析的每一个步骤都有很大的作用，所以该领域在发展发展中是一个值得关注的。此外，鉴于开源项目的本质，在我们完成搜索、收集和编码程序包之后，可以创建更多的程序包，丰富程序包的功能。

### 6.2 展望

为了提高和扩展R程序包，结合本篇研究综述，我们提出了四个建议。第一，使用R进行元分析的局限性之一是，使用命令行分析所需的知识对于一些研究人员来说可能是繁重的，并且经常需要与GUI软件相关的额外培训。因此，我们的第一个建议是程序包的作者可以考虑创建GUI，或者为他们的程序包提供至少一个演示操作，现在已经有一些程序包为用户提供了非常好的示例（如 metafor, metaARRAY, metaMA,和 metaSEM）。尽管现在有很多程序包的作者在CRAN存储库之外的独立的网站上提供演示示例，但这样做并不总是可行的，更何况，把演示示例存储子啊CRAN存储库里也并不难。

1. 根据综述结果显示，R程序包仍然缺少元分析的统计功效估计的功能。具有这个功能的程序包还很少，而且，目前还没有对调节变量进行统计功效分析的程序包。如果程序包中包含此功能，这将对此领域具有很大的作用。
2. 我们建议作者完善元分析程序包制图功能的易操作性，因为目来说使用者细化图形是非常困难的。例如，一个易操作的传统元分析图如森林图、泡沫图、Galbraith plots和漏斗图对使用者都非常有帮助。
3. 程序包的作者应该尽力向初级使用者报告程序包的关键信息，即用户是否能够计算固定效应模型或随机效应模型，这对所有用户在选择使用哪个R程序包是非常重要。通常很难确定程序包的哪些功能可用或者程序包如何实现过程，而这些信息应该包含在所有程序包的文档中。

最后，我们并不对任何一个程序包进行评价，因为所有这些程序包都对社区提供了有益的帮助。通过我们自己的努力，我们认识到寻找一个适当且易于操作的软件包是很困难的，所以这篇综述文章只是为当前和未来的元分析人员提供一个资源。我们相信元分析的实用性，并希望看到用户和分析师社区的成长。通过不断的改进，元分析R程序包有潜力成为该领域强大且有用的工具。

参考文献：

见Polanin, J. R., Hennessy, E. A., & Tanner-Smith, E. E. (2017). A Review of Meta-Analysis Packages in R. Journal of Educational and Behavioral Statistics, 42(2): 206-242.

Table 1 各软件包的特点

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Packages | Version | Title | Effect size | Power | Missing Data | Dependent Effect | Fixed Effect | Random Effect | Moderator Analyses | Publication Bias | Sensitivity Analyses | Creates  Plots | Primary  Function |
| General meta-analysis | | |  |  |  |  |  |  |  |  |  |  |  |
| CAMAN | 0.7 | Finite mixture models and meta-analysis tools |  |  |  |  |  | √ | √ |  |  |  | bivaiate |
| epiR | 0.9-62 | Tools for the analysis of epodemiological data | √ |  | √ |  | √ | √ |  |  |  |  | epi.dsl |
| gmeta | 2.2-3 | Meta-analysis via a unified framework under confidence distribution | √ |  |  | √ | √ | √ |  |  |  | √ | gmeta |
| Mac | 1.1 | Meta-analysis with correlations | √ |  | √ | √ | √ | √ | √ | √ | √ | √ | mareg |
| Mad | 0.8-2 | Meta-analysis with mean differences | √ |  | √ | √ | √ | √ | √ | √ |  | √ | mareg |
| Meta | 4.2-0 | Generalpackage for meta-analysis | √ |  | √ |  | √ | √ | √ | √ |  | √ | metacont |
| metacor | 1.0-2 | Meta-analysis of correlation | √ |  |  |  | √ | √ |  |  |  |  | metacor.DSL |
| metafor | 1.9-5 | Meta-analysis package for R | √ |  | √ | √ | √ | √ | √ | √ | √ | √ | rma |
| metaplus | 0.7-1 | Robust meta-analysis and meta-regression | √ |  |  | √ |  | √ | √ |  | √ | √ | metaplus |
| psychometric | 2.2 | Applied psychometric theory |  |  |  |  | √ |  | √ | √ |  | √ | MetaTable |
| rmeta | 2.16 | Meta-analysis | √ |  |  |  | √ | √ |  | √ | √ | √ | meta.MH |
| Genome meta-analysis | | |  |  |  |  |  |  |  |  |  |  |  |
| EasyStrata | 8.6 | Evaluation of stratified genome wide association meta-analysis results |  |  |  | √ | √ | √ |  |  |  | √ | EasyStrata |
| Gap | 1.1-16 | Genetic analysis package | √ |  |  | √ | √ | √ |  |  |  | √ | metap |
| GeneMeta | 1.40.0 | Meta-analysis for high throughput experiments | √ |  |  |  | √ | √ |  |  | √ | √ | zScores |
| MADAM | 1.2.2 | This package provides some basic methods for meta-analysis |  |  |  |  |  |  |  |  |  | √ | fisher.method |
| MAMA | 2.2.1 | Meta-analysis of MicroArray | √ |  |  |  | √ | √ |  |  |  | √ | ES.GeneMeta |
| metaARRAY | 1.28.0 | Integration of microarray data for meta-analys |  |  |  |  |  | √ |  |  |  |  | poe.em |
| MetABEL | 0.2-0 | Meta-analysis of genome-wide SNP association result | √ |  |  |  |  |  |  |  |  | √ | metagwa.files |
| MetaDE | 1.0.5 | Microarray meta-analysis for differentially expressed gene detection | √ |  | √ |  | √ | √ |  |  | √ | √ | MetaDE.pvalue |
| metaMA | 3.1.2 | Meta-analysis for MicroArrays | √ |  |  |  | √ | √ |  |  | √ |  | directEScombi |
| MetaPCA | 0.1.4 | MetaPCA: Meta-analysis in the dimension reduction of genomic data |  |  | √ | √ |  |  |  |  |  | √ | MetaPCA |
| MetaQC | 0.1.13 | MetaQC: Objective quality control and inclusion/exclusion criteria for genomic meta analysis |  |  | √ | √ |  |  |  |  | √ | √ | MetaQC |
| metaRNASeq | 1.0.2 | Meta-analysis of RNA-seq data | √ |  |  |  | √ |  |  |  | √ |  | fishercomb |
| MetaSKAT | 0.4 | Meta-analysis for SNP-set (Sequence) Kernel Association Test | √ |  | √ |  | √ | √ |  |  |  |  | MetaSKAT\_wZ |
| MultiMeta | 0.1 | Meta-analysis of multivariate genome wide association studies |  |  |  |  |  | √ | √ |  |  | √ | multi\_meta |
| seqMeta | 1.5 | Meta-analysis of region-based tests of rare DNA variants | √ |  |  |  | √ | √ |  |  |  | √ | burdenMeta |
| skatMeta | 1.4.3 | Efficient meta-analysis for the SKAT test | √ |  |  | √ |  | √ |  |  |  | √ | singlesnpMeta |
| Multivariate meta-analysis | | |  |  |  |  |  |  |  |  |  |  |  |
| dosresmeta | 1.3.0 | Performing multivariate dose–dosresmeta response meta-analy | √ |  |  | √ | √ | √ |  |  |  | √ | dosresmeta |
| ecoreg | 0.2 | Ecological regression using aggregate and individual data | √ |  |  |  | √ | √ | √ | √ |  |  | eco |
| metagen | 1 | Inference in meta-analysis and meta-regression | √ |  |  | √ |  | √ | √ |  |  | √ | metareg |
| metaLik | 0.41.0 | Likelihood inference in meta analysis and meta-regression models |  |  |  |  |  | √ | √ |  |  |  | metaLik |
| metaSEM | 0.9-2 | Conducting univariate and multivariate meta-regression with structural equation modeling | √ |  |  | √ | √ | √ | √ | √ |  | √ | meta |
| metatest | 1.0-4 | Fit and test meta-regression model |  |  |  |  |  | √ | √ |  |  |  | metatest |
| mmeta | 2.2 | Multivariate meta-analysis | √ |  |  | √ |  | √ |  |  |  | √ | multipletables |
| mvmeta | 0.4.7 | Multivariate and univariate meta analysis and meta-regression | √ |  | √ | √ | √ | √ | √ |  |  | √ | mvmeta |
| mvtmeta | 1 | Multivariate meta-analysis | √ |  |  | √ | √ | √ |  |  |  |  | mvrmeta\_re |
| robumeta | 1.6 | Robust variance meta-regression |  |  |  | √ |  | √ | √ |  | √ | √ | robu |
| Diagnostic meta-analysis | | |  |  |  |  |  |  |  |  |  |  |  |
| CopulaREMADA | 0.9 | Copula mixed effect models for bivariate and trivariate meta analysis of diagnostic test accuracy studies | √ |  |  |  |  | √ | √ |  |  | √ | CopulaREMADA.norm |
| HSROC | 2.1.8 | Meta-analysis of diagnostic test accuracy when reference test is imperfect | √ |  |  | √ |  | √ |  |  | √ | √ | HSROC |
| mada | 0.5.7 | Meta-analysis of diagnostic accuracy | √ |  |  | √ | √ | √ | √ |  | √ | √ | madauni |
| metamisc | 0.1.1 | Diagnostic and prognostic meta-analysis | √ |  | √ | √ | √ | √ |  |  |  | √ | riley |
| Metatron | 0.1-1 | Meta-analysis for classification data and correction to imperfect reference | √ |  |  | √ |  | √ | √ |  | √ |  | fit.bivar |
| Specific meta-analysis packages | | |  |  |  |  |  |  |  |  |  |  |  |
| compute.es | 0.2-4 | Compute effect sizes | √ |  |  |  |  |  |  |  | √ |  | compute.es |
| CRTSize | 1 | Sample size estimation functions for cluster randomized trials |  | √ |  |  | √ |  |  |  |  |  | fixedMetaAnal |
| exactmeta | 1.0-2 | Exact fixed effect meta-analysis | √ |  |  |  | √ |  |  |  |  |  | meta.exact |
| extfunnel | 1.3 | Additional funnel plot augmentations |  |  |  |  |  |  |  |  |  | √ | extfunnel |
| forestplot | 1.1 | Advanced forest plot using "grid" graphics |  |  |  |  |  |  |  |  |  | √ | forestplot |
| ipdmeta | 2.4 | Tools for subgroup analyses with multiple trial data using aggregate statistic | √ | √ |  |  | √ | √ | √ |  |  | √ | ipdlme |
| metap | 0.6 | Meta-analysis of significance values |  |  |  |  | √ |  |  |  |  | √ | metap |
| metaplot | 1 | Meta-analysis forest plots p drawmeta |  |  |  |  |  |  |  |  |  | √ | drawmeta |
| SCMA | 1.1.1 | Single-case meta-analysis | √ |  |  |  |  |  |  |  |  |  | combine |
| Network meta-analysis | | |  |  |  |  |  |  |  |  |  |  |  |
| gemtc | 0.6-1 | Network meta-analysis using Bayesian method |  |  |  | √ | √ | √ |  |  |  | √ | mtc.model |
| netmeta | 0.7-0 | Network meta-analysis using frequentist methods | √ |  |  | √ | √ | √ |  |  | √ | √ | netmeta |
| pcnetmeta | 2.1 | Patient-centered network meta analysis | √ |  |  | √ | √ | √ |  |  |  | √ | nma.ab |
| Assessment of bias | | |  |  |  |  |  |  |  |  |  |  |  |
| metasens | 0.2-0 | Advanced statistical methods to model and adjust for bias in meta | √ |  |  |  | √ | √ |  | √ | √ | √ | copas |
| PubBias | 1 | Performs simulation study to look for publication bias |  |  |  |  |  |  |  | √ |  | √ | plot\_chase\_observed\_expected |
| SAMURAI | 1.2.1 | Sensitivity analysis of a meta analysis with unpublished but registered analytical investigations | √ |  |  |  | √ | √ |  | √ |  | √ | forestsens |
| selectMeta | 1.0.7 | Estimation of weight functions in meta |  |  |  |  |  | √ |  | √ |  | √ | IyenGreen |
| WeightFunction Model | 1 | The Vevea and Hedges weight function model for publication bias |  |  |  | √ | √ |  | √ |  | √ | √ | n/a |
| Bayesian meta-analysis | | |  |  |  |  |  |  |  |  |  |  |  |
| bamdit | 1.1-1 | Bayesian meta-analysis of diagnostic test data | √ |  |  |  |  | √ |  |  |  | √ | metadiag |
| bspmma | 0.1-1 | Bayesian semiparametric models for meta-analysis |  |  |  |  |  | √ |  |  |  | √ | dirichlet.c |
| Graphical user interface | | |  |  |  |  |  |  |  |  |  |  |  |
| MAVIS | 1.1 | Meta-analysis via Shiny | √ |  | √ | √ | √ | √ | √ | √ | √ | √ |  |
| RcmdrPlugin.MA | 0.0-2 | Graphical user interface for conducting meta-analyses in R | √ |  | √ | √ | √ | √ | √ | √ | √ | √ |  |